

10/070313
JCT rec'd PCT/PTO 05 MAR 2002

Application for
UNITED STATES LETTERS PATENT

of

MASAYUKI ITO

and

YUTAKA YOSHIDA

DATA PROCESSING DEVICE AND DATA PROCESSING SYSTEM

SPECIFICATION

TITLE OF THE INVENTION

DATA PROCESSING DEVICE AND DATA PROCESSING SYSTEM

BACKGROUND OF THE INVENTION

The present invention relates to a data processing device such as a microprocessor or DSP (Digital Signal Processor) having a cache memory and further relates to a data processing system having a burstable memory along with such a data processing device.

In a data processing system using a data processing device such as a microprocessor, a memory preferably for use in attaining high-speed data access or high data transfer performance is a memory that supports burst operations (they are also called burst transfers), typified by a synchronous DRAM (Dynamic Random Access Memory). More specifically, according to this type of memory, an operation control system including circuits such as an internal address counter inside the memory allows reading and writing continuous data at high speed, thereby facilitating the realization of a high-speed, high-performance processing system.

The synchronous DRAM has a mode register. The mode register designates the operation mode thereof. Mode

information including burst length information, which is also called a burst transfer length or block transfer length for the synchronous DRAM, is set in the mode register according to a setting program such as a system initialization program executed after power on reset of a processing system. To set information such as the burst length information, a set period different from the burst operations is required. That is, it relatively takes time to set the mode information in the mode register. On this account, the burst length information once set in accordance with power on reset is not changed later in general.

Here, in the burst operations, when the burst length is set greater, an amount of data that can be transferred by access at one time to the synchronous DRAM becomes large, and thus the high data transfer performance can be obtained in the case of transferring a mass amount of data. However, when an amount of data smaller than the set burst length is transferred, unnecessary data transfer cycles are increased and the data transfer performance is deteriorated because block transfers are conducted according to the set burst length.

In the memory that supports burst transfers, typified by the synchronous DRAM, in burst transfers a wraparound function is supported in which transfer can be started from data at an arbitrary address between the boundaries of burst transfers. As for location addresses between the boundaries,

0205060708090A0B0C0D0E0F

the first location of data to be accessed is designated from outside, and the subsequent location addresses are created by an internal counter such as a column address counter inside the memory. The SDRAM is considered to have an access unit of four bytes and a burst length of 16 bytes, for example, in which a column address in a byte unit is preset in the column address counter, and it sequentially undergoes counter operations for three times from the least significant bit to the fourth bit, starting from the preset address, whereby continuous access operations are performed. For example, when data locations in a four-byte unit between the boundaries of burst operations are set to an address of $N + 00$ (hereafter, it is denoted as @00), @04, @08 and @12, data access is performed by wraparound in order of @08, @12, @00, and @04, where the first location of data accessed in burst operations is set to @08.

Utilizing the wraparound function can acquire data required by a CPU from an external memory at the start of burst transfers. Accordingly, in a cache mishit, for example, the number of cycles that the CPU waits for data can be reduced.

The memory that supports burst access, typified by the synchronous DRAM, gives the characteristics according to the burst length as described above. Thus, desirably, the type of memory can be adapted to both demands of setting the burst length greater and setting it smaller. Then, the inventors

investigated the effectiveness of a control method of utilizing both of memories performing wraparound at a different burst length, such as a synchronous DRAM having a burst length of 32 bytes and a synchronous DRAM having a burst length of 16 bytes.

The inventors revealed the following items according to the investigation. More specifically, the order of data returned from the synchronous DRAM in the wraparound operation is varied in the case where the synchronous DRAM having a burst length of 16 bytes undergoes burst operations by wraparound starting from the eighth byte in the column address location of 16 bytes and in the case where the synchronous DRAM having a burst length of 32 bytes undergoes burst operations by wraparound starting from the eighth byte in the column address location of 32 bytes. Thus, the need occurs in which a memory control circuit for controlling the synchronous DRAM is configured to have a scheme of recognizing or eliminating a mismatch of data arrangement due to the burst length difference.

The two following techniques can be named as the scheme to do so, for example.

A first technique is that two sets of data of 16 bytes returned from the memory having a block transfer length of 16 bytes are conformed to the same data order as the data order in the wraparound operation of a burst length of 32 bytes. In

this case, a buffer memory and an aligner for sorting data are set along with the memory control circuit. The memory control circuit performs the control operation in which in burst read operations, for example, the memory control circuit temporarily buffers data outputted from the memory in the buffer memory, and then it uses the aligner to sort data for data output so that two sets of data having undergone wraparound in 16 bytes are conformed to the data order obtained by the wraparound operation of 32 bytes. In the case of the first technique, extra waiting time is needed in data buffering for sorting data.

A second technique is that a mismatch of the order of returning data is prevented from occurring by establishing a restriction where a start memory access address has the boundary fixing of 16 bytes. Accordingly, the data order when burst operations are performed twice in a burst length of 16 bytes is conformed to the data order when a burst operation is performed once in a burst length of 32 bytes.

However, the inventors took notice that the both techniques described above are accompanied by the deteriorated data processing performance. Hereafter, it will be described.

Here, the microprocessor is considered to have a CPU, a cache memory and a memory control circuit for having access to external memories including an external synchronous DRAM. Now, consider the case where a cache line length of the cache

memory is 32 bytes, a cache mishit is made when the CPU starts access at an address of $N + 08$ (N is a multiple of 32) in a memory, the external memories are accessed according to that, cache lines are filled, and then the CPU continuously requests data at sequential addresses of $N + 12$, $N + 16$, $N + 20$, $N + 24$, and $N + 28$. Additionally, the access to such a sequential addresses can be considered to be a natural example frequently generated in the case of instruction access or in processing data arranged in continuous areas. Hereafter, data at an address of $N + 08$ is denoted by @08, and data at an address of $N + 12$ is denoted by @12.

Under the conditions, in the first technique, the data order obtained from the memory according to the burst operations of a block transfer length of 16 bytes is considered to be in order of @08, @12, @00, @04, @24, @28, @16, and @20, for example. In sorting it into a data order as similar to the data order in the case of the burst operations of a burst length of 32 bytes and returning it to the cache memory, some penalty cycles are generated, which cause the bus performance or data processing performance of the CPU to be deteriorated. More specifically, in the case of that data order, the data order corresponding to the burst operations of 32 bytes is in order of @08, @12, @16, @20, @24, @28, @00, and @04. In sorting the data in that order and retuning it to the cache memory, data @16 needed to return to third position corresponding to

the burst operations of 32 bytes arrives from the memory in the seventh place in the burst operations of 16 bytes, which will generate at least four penalty cycles.

In the second technique, both the data order from the external memories and the data order to the cache corresponding to the case where the CPU first needs data @08 are in order of @00, @04, @08, @12, @16, @20, @24, and @28. That is, data @08 required first by the CPU is to be the third data. In response to this, the CPU will wait the arrival of the first data needed by the CPU for at least two cycles. In this manner, the deteriorated performance of the CPU occurs in the second technique as well.

Additionally, the ninth page and Table 2 of MICROPRESSOR REPORT VOL. 10, NO. 2, FEBRUARY 12, 1996 describe that the order of the cache fill is controlled so as to put a missed word of the cache memory first in a 32-byte microprocessor. However, it does not take notice of the relationship between the burst length of the synchronous DRAM and the data order of cache fill.

A purpose of the invention is to provide a data processing device capable of shortening waiting time of a CPU until acquiring data associated with cache miss and of contributing to the enhanced data processing performance, even using a memory burstable in a size shorter than a cache line length of a cache memory and having a wraparound function, and to

further provide a data processing system.

Another purpose of the invention is to provide a data processing device having fewer penalty cycles in memory access and capable of enhancing the bus performance and the CPU performance, even connecting and utilizing a plurality of memories having a wraparound function and a different burst length, and to further provide a data processing system.

Still another purpose of the invention is to provide a data processing device capable of responding to various connection configurations and utilization forms of burstable memories having a wraparound function.

The above and other purposes and the novel features of the invention will be apparent from the following description of the specification and accompanying drawings.

SUMMARY OF THE INVENTION

[1] The data processing device has a CPU, a cache memory accessible by the CPU, a cache control part for controlling the cache memory, and a memory control part accessible to memories in response to a cache mishit of the cache memory. When the memory control part has access to a burstable memory in response to a cache mishit, it forms first information for indicating a burst length of the memory to a cache line length of the cache memory, and it can control a single or plurality of burst operations necessary to obtain a data length meeting

the cache line length according to the first information. The cache control part can control the cache fill operation of filing data acquired in the single or plurality of burst operations in the cache memory by wraparound according to the first information.

According to the scheme described above, the first information reveals the burst length of the memory to be accessed to the cache line length, the number of burst operations corresponding to the memory to be accessed is controlled according to this, and block data corresponding to the cache line length can be acquired from the memory in the burst operations. The cache control part allows the block data, which is being transferred to the cache memory in the wraparound operation, to be filled in the cache, as meeting the acquired block data to the burst length learned from the first information. On this account, data outputted from the memory does not need to be sorted by an aligner. Additionally, the restriction where the top boundary of the data block to be the object for burst operations is fixed to a start access address does not need to be provided. Thus, the data processing device can shorten waiting time of the CPU until acquiring data in a cache miss, even using the memory burstable in a size shorter than the cache line length of the cache memory and having a wraparound function, and therefore it can contribute to the enhanced data processing performance.

In the cache fill operation, the cache control part can be configured to receive the inputs of address information in a cache miss, the first information and a synchronization signal synchronous with a delimiter of data obtained in the burst operations by the memory control part, to perform the wraparound control starting from the address information in the range of the burst length meant by the first information, and to create cache fill addresses for determining the data order of cache fill in synchronism with the synchronization signal.

Accordingly, the cache fill operation can proceed as following the operation where the memory control part sequentially reads data out of the memory in burst operations in response to a cache mishit. Regardless of a burst length of the memory, the high-speed cache fill operation can be assured.

The memory control part can be configured to control burst operations in which in having access to a memory in a plurality of burst operations in response to a cache mishit, the memory control part controls a burst operation by wraparound starting from a data location at an address in the cache mishit in the first burst operation, and it controls burst operations starting from the top boundary of the data block defined by the burst length in the subsequent burst operations.

In view of the fact that a program or certain data is

often located at sequential addresses, as for access other than the first access among a plurality of access in burst operations, when the memory is accessed from the top boundary defined by burst length, data to be accessed first by the CPU is to reach the cache memory or CPU first in sequential data access, as described above. Therefore, it is useful to enhance the data processing performance.

According to the data processing device described above, when a burst length (16 bytes, for example) relatively shorter than the cache line length (32 bytes, for example) of the cache memory is set in the memory, a plurality of data blocks outputted by wraparound from the memory performing burst access operations can be joined and filled in the cache memory. Furthermore, in the case where write through is adopted as one process for a cache write hit of the cache memory, when write data is written into the memory from a write through buffer shorter than the cache line length (eight byte, for example), there are fewer unnecessary data transfer cycles because of the relatively shorter burst length. As for eight bytes in the latter half of the burst access operation at this time, the actual data write operation can be suppressed by data masking.

Accordingly, the data processing device can suppress unnecessary cycles to the minimum in transferring relatively small data of eight bytes without deteriorating the performance

of efficiently transferring high-capacity data of 32 bytes to the cache memory, and thus it can enhance the data processing performance.

[2] The data processing system has a data processing device having a CPU and a cache memory, and a memory connected to the data processing device, the memory is burstable and configures a main memory for the cache memory. The memory may be single or plural. The burst length of each memory may be different or the same. The cache memory has a cache line length of L bytes. The memory is burstable by wraparound in the range of a burst length in bytes, where L times one over two to the n-th power (n is a natural number). At this time, the data processing device performs control in which it forms first information for indicating the burst length of the memory to the cache line length of the cache memory in response to a cache mishit of the cache memory, it allows the memory into burst operations for a single or plurality of times necessary to obtain a data length meeting the cache line length according to the first information, it joins a plurality of block transfer data thereby acquired, and it returns data of L bytes to the cache memory.

According to the data processing system, when a burst length (16 bytes, for example) relatively shorter than the cache line length (32 bytes, for example) of the cache memory is set in a first memory, a plurality of block data outputted

from the burstable memory by wraparound can be joined and filled in the cache memory. When the data processing system includes a second memory set up with the same burst length as the cache line length, it allows the cache fill operation responding to the burst length of the second memory in processing a cache mishit for the second memory.

Additionally, in the case where write through is adopted as a process for a cache write hit of the cache memory, when write data is written into the first memory from a write through buffer shorter than the cache line length (eight bytes, for example), there are fewer unnecessary data transfer cycles because of the relatively shorter burst length. As for eight bytes in the latter half of the burst access operation at this time, the actual data write operation can be suppressed by data masking. When the second memory set up with the same burst length as the cache line length is the object to be written by write through, unnecessary cycles are increased as compared with the first memory, even performing write mask. Even so, in a state that the second memory is temporarily removed from the object for cache, it is possible to increase an amount of data accessible or transferred at one time.

Accordingly, the data processing system can suppress unnecessary cycles as much as possible in transferring relatively small data of eight bytes without deteriorating the performance of efficiently transferring high-capacity data of

32 bytes to the cache memory, it can realize various connection configurations or utilization forms of a plurality of memories having a different burst length, and thus the data processing performance can be enhanced in the data processing system.

Also in the data processing system, the data processing device can be configured to perform control of filling data, which is being transferred in the wraparound operation, in the cache memory according to the first information, the data has been acquired in the single or plurality of burst operations. Also at this time, in the cache fill operation, the data processing device can be configured in which it creates a synchronization signal synchronous with a delimiter of data obtained from the memory in the burst operations, it controls wraparound starting from the address information in the range of a burst length meant by the first information, and it creates cache fill addresses for determining the data order of cache fill in synchronism with the synchronization signal. Furthermore, the data processing device can be configured in which in having access to a memory in burst operations for a plurality of times in response to a cache mishit, it controls a burst operation by wraparound starting from the data location at the address in the cache mishit in the first burst operation, and it controls burst operations starting from the top boundary of the data block defined by the burst length in the subsequent burst operations.

BRIEF DESCRIPTION OF THE DRAWINGS

The teachings of the invention can be readily understood by considering the following detailed description in conjunction with the accompanying drawings, in which:

Fig. 1 depicts a block diagram illustrating one example of a data processing system in the invention;

Fig. 2 depicts a block diagram illustrating one example of a detailed block transfer length determining part;

Fig. 3 depicts a block diagram illustrating one example of an external memory address generating circuit;

Fig. 4 depicts an illustration exemplifying the address generation rule of the generation logic for subsequent access addresses;

Fig. 5 depicts a timing chart exemplifying burst operations for a synchronous DRAM set up with a burst length of 32 bytes;

Fig. 6 depicts a timing chart exemplifying burst operations for a synchronous DRAM set up with a burst length of 16 bytes;

Fig. 7 depicts a block diagram illustrating one example of the logic configuration for generating a cache access address and a memory access address in a cache control part;

Fig. 8 depicts an illustration exemplifying the address generation logic of a cache fill address generating circuit;

Fig. 9 depicts a timing chart illustrating the cache fill operation by a microprocessor shown in Fig. 1, also including comparative examples; and

Fig. 10 depicts a block diagram illustrating another example of the data processing system in the invention.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

Fig. 1 depicts one example of the data processing system in the invention. The data processing system depicted in the same drawing is typically provided with a microprocessor 1, which is one example of the data processing device in the invention, a synchronous DRAM 2, which is one example of a burstable external memory, and a ROM (Read Only Memory) 14. It may be provided with other peripheral circuits other than the synchronous DRAM 2 and the ROM 14.

Although the microprocessor 1 is not defined particularly, it has a CPU 3, a cache memory 4, a cache control part 5, and a memory control part 6. They are formed on one semiconductor substrate (semiconductor chip), for example. Although data buses 8, 9 and 10 are not defined particularly, they are considered to be four bytes (32 bits).

The CPU 3 has a control part and an execution part, not shown in the drawing. The execution part has a general purpose register file and a computing unit, for example. The control part decodes fetched commands and it controls the operation

of the execution part.

The cache memory 4 has a so-called data array. The data array is configured of an SRAM (Static Random Access Memory), for example, and it has a matrix of memory cells. A selection terminal of the memory cells is connected to a word line at every row, for example, and a data input/output terminal of the memory cells is connected to a complementary bit line at every column. In the data array, the word line is selected by an index address given from the cache control part 5. A unit at every row selected by the index address is a cache line in the data array. Although the cache line is not defined particularly, it has a cache line length of 32 bytes. For the selected cache line, a long word selection signal given from the cache control part 5 performs the selection of four bytes. The index address and the long word selection signal are denoted by a cache access address signal 7 in the drawing.

The cache control part 5 has a so-called address array and a cache control logic. The address array is also configured of the SRAM as similar to the data array. The address array has tag fields for each of the cache lines one on one. The tag field holds valid bits for indicating the effectiveness of the corresponding cache line tag or the cache line. The tag field is also selected by the index address. The cache control logic determines whether a cache hit or cache mishit, and it controls cache fill when a cache mishit.

The memory control part 6 performs bus control for having access to the synchronous DRAM 2 or ROM 14 according to the instruction of the CPU 3 or cache control part 5. The memory control part 6 is connected to the synchronous DRAM 2, typically shown, through the external bus 10 and an external address bus 13. Control buses for transmitting control signals such as a strobe signal for access to the external busses or external memories are omitted from the drawing. The memory control part 6 can be understood as a so-called bus state controller, or a part of a memory controller contained therein.

A part of an effective address 11 outputted by the CPU 3 is set to be the index address. The cache control logic compares the tag of the tag field indexed in the address array with a tag address contained in a part of the effective address 11. It is considered to be a cache hit when they match each other, whereas it is considered to be a cache mishit when they do not match each other.

In the read access by the CPU 3, when it is a cache hit (cache read hit), data of four bytes corresponding to the indexed cache line is fed to the CPU 3 through the data bus 8. When it is a cache mishit (cache read mishit) in the read access, the cache control part 5 generates a memory access address 12 and it gives a memory access request MREQ to the memory control part 6 along with the memory access address 12. Accordingly, the memory control part 6 reads data of one cache

line from the synchronous DRAM 2, for example, and it feeds the read data to the cache memory 4 through the data bus 9. In synchronism with this, the cache control part 5 generates the cache access address 7 and it fills the data in a required cache line. Additionally, the cache control part 5 stores a tag corresponding to the data of the cache line in a tag field corresponding to the cache line. At this time, data in a cache miss is given to the CPU 3 through the data bus 8.

In the write access by the CPU 3, when it is a cache hit (cache write hit), the CPU 3 feeds write data to four bytes of the corresponding cache line indexed through the data bus 8. When it is a cache mishit (cache write mishit) in the write access, the cache control part 5 generates the memory access address 12 and it gives the memory access request MREQ to the memory control part 6. The memory control part 6 reads data of one cache line from the synchronous DRAM 2, for example, according to the memory access address 12, and it gives the read data to the cache memory 4 through the data bus 9. In synchronism with this, the cache control part 5 fills the data in a cache line, and it stores a tag corresponding to data of the cache line in the tag field corresponding to the cache line.

As for a scheme of maintaining the match between data held by the cache memory 4 and data stored in the external memory such as the synchronous DRAM 2, a method of write through is adopted, for example. That is, the cache memory 4 has a write

through buffer for holding write data when a cache write hit, not shown in the drawing. When a cache write hit, the cache control part 5 writes write data in the cache write hit into the cache memory 4, and then it gives an instruction of writing into the corresponding address of the corresponding synchronous DRAM 2 to the memory control part 6. Accordingly, the memory control part 6 controls writing data held in the write through buffer into the synchronous DRAM 2.

The synchronous DRAM 2 has a memory cell array having a matrix of dynamic memory cells in which the information storage format is performed dynamically through a storage capacitor as similar to a DRAM, and the refreshment of storage information is also needed. A huge difference from the DRAM is in that the synchronous DRAM 2 is operated synchronously with an external clock signal and it is burstable by wraparound. For example, it has a column address counter for latching a column address signal fed from outside in which it sequentially updates column addresses by the column address counter and it can efficiently perform the continuous data access operation, starting from a preset value of the column address counter as maintaining the word line selecting state by row addresses. The number of continuous data access is called a burst length, and the column address counter undergoes counter operations only for the number of times defined by the burst address. For instance, in a synchronous DRAM having an access unit of four

bytes and a burst length of 16 bytes, the continuous access operation can be performed in which the column address in a byte unit is preset in the column address counter and the column address counter undergoes counter operations for three times from the least significant bit to the fourth bit, starting from the preset address. Therefore, when the base point of access in a four-byte unit is not at the boundary of the location in the column data location of 16 bytes, the counter address by the column address counter is returned to the boundary of the previous column data location of 16 bytes from the subsequent column data location of 16 bytes in the midway for address counting. That is, the access order of the burst operations is performed by wraparound within the column data location of 16 bytes.

The burst length is set in a mode register of the synchronous DRAM 2. For example, a part of memory control data 15 set in the memory control part 6 from the CPU 3 in the power on reset processing is also initialized in the mode register by the CPU 3 as data for indicating a burst length. Although a burst length is not defined particularly, it can be set selecting from 16 or 32 bytes. The operation of the synchronous DRAM 2 is instructed according to the states of signals such as row address strobe (RAS), column address strobe (CAS), write enable (WE), data mask (DM), and data strobe (DQS). The memory control part 6 generates the signals. Commands are defined according

to each of the specific states of the signals, and the synchronous DRAM 2 performs operations according to the instruction of the command. For example, an active command accompanied by the row address signal instructs a word line selection operation. A read command accompanied by the column address signal instructs a read operation for the memory cells of the word line already activated. Additionally, a write command accompanied by the column address signal instructs a write operation for the memory cells of the word line already activated. The read operation and the write operation are performed in wraparound-capable burst access in a burst length set in the mode register. In the write operation, in an access cycle where the data masking (DM) signal is set enable, only the access cycle is consumed and the actual data write is suppressed.

Next, the memory control and the cache fill operation corresponding to the burst length of the synchronous DRAM 2 will be described.

First, the summary thereof will be described in accordance with Fig. 1. A block transfer length determining part 20 and an external memory address generating part 30 are typically depicted in the memory control part 6. The block transfer length determining part 20 forms wraparound information WRPA that is the first information for indicating the burst length of the synchronous DRAM 2 to the cache line

2025 RELEASE UNDER E.O. 14176

length (32 bytes) of the cache memory 4 in having access to the synchronous DRAM 2 in response to the memory access request MREQ from the cache control part 5 due to a cache mishit. The external memory address generating part 30 controls a single or plurality of burst operations necessary to obtain the data length meeting the cache line length and it burst-reads data out of the synchronous DRAM 2, according to the wraparound information WRPA. It may activate burst access twice when the burst length is 16 bytes, whereas it may activate burst access once when the burst length is 32 bytes. The top of burst read is data of an address in a cache miss. The cache control part 5 generates cache fill addresses for writing data of 32 bytes read by the memory control part 6 in burst read into the cache memory 4 by wraparound at every four bytes. The data block of one wraparound operation at this time is matched with the burst length of the synchronous DRAM 2. It undergoes wraparound operations at every range of a 16-byte address when the burst length is 16 bytes, whereas it undergoes wraparound operations at every range of a 32-byte address when the burst length is 32 bytes. The cache fill address in the wraparound operations is the index address and the long word selection signal 7. The long word selection signal is synchronous with a data ready signal DRDY for indicating a delimiter of data that the memory control part 6 reads in burst read and outputs to the data bus 9 at every four bytes.

Fig. 2 depicts a detailed example of the block transfer length determining part 20. The block transfer length determining part 20 has an access request determining circuit 22, memory control registers 23, and a block transfer length determining circuit 24. In the memory control registers 23, the CPU 3 initializes a data bus width for an external address area of the microprocessor 1, the number of access cycles, external memory information 15 such as a burst length. Burst length data for indicating a burst length having been set in the synchronous DRAM 2 by the CPU 3 is also set in the memory control registers 23.

The access request determining circuit 22 receives the inputs of the memory access request MREQ and the memory access address 12 from the cache control part 5, and it activates a detection signal 25 when an object to be accessed is the synchronous DRAM 2. Furthermore, the access request determining circuit 22 detects a memory access request by the memory access request MREQ, it decodes the memory access address 12, it selects an area according to the area to be accessed, and then it generates an access area selection signal (not shown). This access area selection signal is utilized as a chip selection signal or memory enable signal for memories, for example.

The block transfer length determining circuit 24 receives the inputs of the detection signal 25 and information

26 for a burst length of the synchronous DRAM 2 set in the memory control registers 23 and it outputs the wraparound information WRPA. Here, the burst length of the synchronous DRAM 2 is 16 or 32 bytes and the cache line length of the cache memory 4 is 32 bytes. Thus, although the wraparound information WRPA is not defined particularly, it is information of one bit; the logical value "0" means a burst length of 16 bytes, and the logical value "1" means a burst length of 32 bytes, for example.

Fig. 3 depicts one example of the external memory address generating part 30. The external memory address generating part 30 has an address buffer 31, a subsequent access address generating circuit 32, and a selector 33. The external memory address generating part 30 receives the memory access address 12 from the cache control part 5 and then it keeps the memory access address 12 in the address buffer 31. Subsequently, it allows the selector 33 to select the address kept in the address buffer 31 and to output the address to the address bus 13 as an external memory address 16. When the area selection by the request determining circuit 22 at this time is the synchronous DRAM 2, the synchronous DRAM 2 is selected as a chip, and a read or write command is fed through a synchronous DRAM control logic inside the memory control part 6, not shown. Thus, the synchronous DRAM 2 is allowed into burst operations. When the wraparound information WRPA is the logical value "1", the burst operation is finished for one time. When the wraparound

information WRPA is the logical value "0", the subsequent access address generating circuit 32 performs +16 to an address (byte address) of the address buffer 31, for example, according to the after-mentioned address generation logic and it generates the first address in the next burst operation, in order to allow the synchronous DRAM 2 into burst operations twice. The details of the generation logic for the subsequent access addresses will be described later. In the second-time burst operation, the selector 33 selects the output of the subsequent access address generating circuit 32 to feed it to the synchronous DRAM 2.

Fig. 4 exemplifies the rule of the generation logic for the subsequent access addresses. Here, supposing that the cache line length is 32 bytes, the burst length of the synchronous DRAM is 32 or 16 bytes, and the data bus width is four bytes, where N is a multiple of 32, data of four byte from an address of N is denoted by D1, and data of four byte from an address of N + 4 is denoted by D2.

In Fig. 4, a first access address means a start address of the first-time burst operation, and a second access address means a start address of the second-time burst operation necessary when the burst length is 16 bytes. The access address at the second time is not set to a value where the access address at the first time is uniformly added with 16 bytes. In the case where the burst length is 16 bytes, the access address

at the second time is set to an address of $N + 16$ when the first access address is addresses of $N + 4$, $N + 8$, and $N + 12$. Therefore, the data output is in order of addresses at the second-time burst access. In view of the fact that a program or certain data is often located at sequential addresses, as described above, as for access except the first access among a plurality of access in burst operations, when the memory is accessed from the top of the boundary defined by the burst length, data to be accessed first by the CPU 3 arrives at the cache memory 4 or CPU 3 first in the continuous data access, and thus it is useful to enhance the data processing performance. In conformity to this, the second-time access address is set to an address of $N + 0$ when the first access address is also addresses of $N + 20$, $N + 24$, and $N + 28$.

Fig. 5 exemplifies a timing chart of the burst operations for the synchronous DRAM 2 set up with a burst length of 32 bytes. Here, a start transfer address given by the external memory address generating part 30 is an address of $N + 8$, and the wraparound information means 32 bytes. In this case, the synchronous DRAM does of course not need the second-time burst access. In Fig. 5, it should be understood that a bank active command, not shown in the drawing, has been issued prior to the read command and the word selection operation has already been completed. As apparent from Fig. 5, a burst read of 32 bytes is performed by wraparound in order of D3, D4, D5, D6,

D7, D8, D1, and D2.

Fig. 6 exemplifies a timing chart of the burst operations for the synchronous DRAM 2 set up with a burst length of 16 bytes. Here, it should be understood that the first start transfer address given by the external memory address generating part 30 is an address of $N + 8$, the start transfer address for the burst operation at the second-time is also set to an address of $N + 16$ in accordance with Fig. 4, the bank active command, not shown in the drawing, has first been issued prior to the read command, and the word selection operation has already been completed. As apparent from the logic shown in Fig. 4, a burst read is performed by wraparound in order of D3, D4, D1, and D2 in the first-time burst operation. A burst read is performed in order of D5, D6, D7, and D8 from the top of the data block in the second-time burst operation.

Fig. 7 depicts one example of the logic configuration for generating cache access addresses and memory access addresses in the cache control part 5. The cache control part 5 has an address buffer 40, a memory access address generating circuit 41, a cache fill address generating circuit 42, and a selector 43. The cache control part 5 receives an effective address 11 from the CPU 3 and it keeps the effective address 11 in the address buffer 40. Then, it allows the selector 43 to select the address kept in the address buffer 30 and to feed the address to the cache memory 4 as the cache access address

7. When a cache entry of the cache line corresponding to the cache access address is a cache mishit, the memory access address generating circuit 41 generates the memory access address 12 in the cache mishit in response to this. The access control of the synchronous DRAM 2 by the memory control part 6 using the memory access address 12 is as described above.

The cache fill address generating circuit 42 generates the cache fill addresses for writing data of 32 bytes that the memory control part 6 has read from the synchronous DRAM 2 in burst read into the cache memory 4 by wraparound at every four bytes. The cache fill address generating circuit 42 at this time receives the input of the wraparound information WRPA so as to allow the data block in wraparound operations to meet the burst length of the synchronous DRAM 2, and it allows wraparound operations in the range of every address of 16 bytes when the burst length is 16 bytes, whereas it allows wraparound operations in the range of every address of 32 bytes when the burst length is 32 bytes. The first address of the cache fill addresses in wraparound operations is an address in a cache miss held in the address buffer 40. The cache fill addresses in wraparound operations are the index address and the long word selection signal 7 as described above. When the memory control part 6 outputs data having been read out of the synchronous DRAM 2 in burst read to the data bus 9 at every four bytes, it outputs a data ready signal DRDY for indicating

the delimiter of the data. The cache fill address generating circuit 42 sequentially increments the first address of the cache fill addresses by +4 in synchronism with the data ready signal DRDY.

Fig. 8 exemplifies the address generation logic of the cache fill address generating circuit 42. The first address of the cache fill addresses is determined by the effective address in a cache mishit. Thus, the cache fill addresses of a burst length of 32 bytes and 16 bytes and data corresponding thereto are shown at every effective address in a cache mishit in pairs in Fig. 8.

In Fig. 8, N is set to a multiple of 32, data at an address of N is called D1, data at an address of $N + 4$ is D2, and data at an address of $N + 28$ is D8. At this time, when the effective address from the CPU 3 is an address of $N + 8$ and the wraparound information means 16 bytes, the data order returned from the synchronous DRAM 2 is in order of D3, D4, D1, D2, D5, D6, D7, and D8. To fill them in the proper locations of cache, the cache fill address generating circuit 42 generates the cache access address 7 so as to sort the cache fill addresses in order of $N + 8$, $N + 12$, N , $N + 4$, $N + 16$, $N + 20$, $N + 24$, and $N + 28$, in accordance with the data ready signal DRDY issued by the memory control part 6 in synchronism with data change.

Fig. 9 depicts the above-described cache fill operation by the microprocessor 1, also including comparative examples.

In the related art, when an address in a cache mishit is set to an address of $N + 8$, data read out of a synchronous DRAM of a burst length of 32 bytes in burst operations is in order of @08, @12, @16, @20, @24, @28, @00, and @04, as shown in (A) in Fig. 9. Data read out of a synchronous DRAM of a burst length of 16 bytes in twice burst operations is in order of @08, @12, @00, @04, @24, @28, @16, and @20, as shown in (B) in Fig. 9. In this manner, the data order read out of the synchronous DRAM is varied from a burst length. In the related art in which a data aligner for conforming the data order to the data order in burst operations of 32 bytes is provided before cache fill so as not to perform cache fill with this mismatch, four penalty cycles in data sorting are generated, as shown in (C) in Fig. 9, and thus the bus performance is deteriorated. On the other hand, in the related art in which the start access address has the restriction of boundary fixing, as shown in (D) in Fig. 9, since the first data is fixed to an address of $N + 0$ for boundary fixing in spite of having been generated a cache mishit at an address of $N + 8$, data to be required first by the CPU 3 cannot be fed in the first place to generate two cycles of penalty, thereby causing the CPU performance to be deteriorated. However, in the microprocessor 1, the process of buffering and sorting burst read data is not performed, and the memory control part generates cache fill addresses corresponding to the burst length of burst read data according to the wraparound

information, as shown (E) in Fig. 9. Therefore, the penalty is not generated as shown in (C) in Fig. 9. Additionally, as apparent from (E) shown in Fig. 9, the start address of the burst access does not have the restriction of boundary fixing, and thus the penalty is also not generated as shown in (D) in Fig. 9. Furthermore, the first address of the second-time burst access is set to the top of the boundary according to the logic shown in Fig. 4, as (E) shown in Fig. 9. Therefore, when the CPU 3 requires continuous data, the request can be responded promptly.

The operation will be described more specifically, with reference to Fig. 9. In the case where the CPU 3 starts access from an address of $N + 08$, memory access due to a cache miss is generated and the CPU 3 continuously requires data of sequential addresses of $N + 12$, $N + 16$, $N + 20$, $N + 24$, and $N + 28$, the data order from the synchronous DRAM is in order of @08, @12, @00, @04, @24, @28, @16, and @20 in the scheme of (C) shown in Fig. 9. Data @16 that the CPU 3 requires third arrives in the seventh place from the memory, and thus at least four penalty cycles are generated. However, when using the scheme of (E) shown in Fig. 9, the data order from the synchronous DRAM 2 is in order of @08, @12, @00, @04, @16, @20, @24, and @28. Data @16 that the CPU 3 requires third is allowed to arrive in the fifth place from the synchronous DRAM 2, and the penalty cycle can be reduced to two cycles. Thus, the

enhanced data processing performance by the CPU 3 can be realized. Such the access to the sequential addresses is generated with great frequency in the case of instruction access or continuous data processing. Therefore, the great effect can be obtained in enhancing the data processing efficiency.

Moreover, the cache control part 5 receives the wraparound information WRPA along with data from the synchronous DRAM 2, whereby access can also be started from the location other than the boundary top of the data block defined by burst length and the data processing performance by the CPU can be enhanced. More specifically, in (D) shown in Fig. 9, data required first by the CPU is data @08, but the data order returned from the memory has the restriction of starting from the top of the memory block. Thus, memory access from data @08 cannot be performed, the start burst transfer address is turned to be an address of $N + 0$, and the data order retuned from the memory is in order of @00, @04, @08, @12, @16, @20, @24, and @28. Consequently, data @08 is in the third place, and the CPU has to wait the arrival of the first data for two cycles, thereby causing the deteriorated data processing performance of the CPU. However, when using the control scheme typified by (E) shown in Fig. 9, the start burst transfer address can be turned to be an address of $N + 8$, the data order from the memory is in order of @08, @12, @00, @04, @16, @20,

@24, and @28. Data @08 required first by the CPU 3 is allowed to arrive in the first place from the synchronous DRAM 2, the penalty cycles of the CPU 3 can be reduced to two cycles, and therefore the enhanced data processing performance of the CPU 3 can be realized.

Besides, the process of (E) shown in Fig. 9 adapts the generation logic for the cache fill addresses shown in Fig. 8. Thus, as shown the same drawing, when data @010 required third by the CPU 3 is obtained, four penalty cycles generated in (C) shown in Fig. 9 can be suppressed to two cycles. Also in this point, the data processing performance by the CPU 3 can be enhanced.

The operation and effects obtained by the above-described microprocessor 1 will be organized and described.

In the microprocessor 1, the memory control part 6 learns the burst length of the memory to be accessed (synchronous DRAM 2) to the cache line length according to the wraparound information WRPA, it controls the number of burst operations corresponding to the memory 2 to be accessed according to this, and it can obtain the block data corresponding to the cache line length from the synchronous DRAM 2 in burst operations. The cache control part 5 allows the acquired block data to be filled in the cache memory 4 by wraparound as the block data meets the burst length learned from the wraparound information WRPA. Thus, data outputted from the synchronous DRAM 2 does

not need to be sorted by the aligner, and the restriction where the top boundary of the data block that undergoes burst operations is fixed to the start access address does not need to be provided. Accordingly, even when using the memory burstable in a size shorter than the cache line length of the cache memory and having a wraparound function, the microprocessor 1 can shorten waiting time of the CPU until acquiring data in a cache miss, and therefore it can contribute to the enhanced data processing performance.

Additionally, the cache control part 5 can advance the cache fill operation as following the operation where the memory control part 6 sequentially reads data out of the synchronous DRAM 2 in burst operations in response to a cache mishit. Thus, the high-speed cache fill operation can be assured.

In view of the fact that a program or certain data is often located at sequential addresses, as described above, as for access except the first access among a plurality of access in burst operations, data to be accessed first by the CPU is allowed to arrive at the cache memory or CPU first in continuous data access when the memory is accessed from the top boundary defined by burst length. Therefore, it is useful to enhance the data processing performance.

The data processing system exemplified in Fig. 1 has one synchronous DRAM 2 connected to the microprocessor 1. When

25 20 20 20 20 20 20

a burst length (16 bytes, for example) relatively shorter than the cache line length (32 bytes, for example) of the cache memory 4 is set in the synchronous DRAM 2, a plurality of block data outputted by wraparound from the synchronous DRAM 2 performing the burst access operation can be joined and filled in the cache memory 4. Furthermore, in the case where write through is adopted as one process for the cache write hit of the cache memory 4, when write data is written into the synchronous DRAM 2 from the write thorough buffer shorter than the cache line length (eight bytes, for example), there are fewer unnecessary data transfer cycles because of the relatively shorter burst length. For eight bytes in the latter half of the burst access operation at this time, the data masking signal DM masks data, whereby the actual data write operation can be suppressed.

Accordingly, the data processing system does not impair the performance of efficiently transferring high capacity data of 32 bytes to the cache memory 4, it can suppress unnecessary cycles to the minimum in transferring relatively small data such as the data write through operation of the write through buffer, and thus it can enhance the data processing performance.

Fig. 10 depicts another example of the data processing device. The data processing system shown in the same drawing is provided with memories burstable by wraparound, such as two

00000000000000000000000000000000

synchronous DRAMs 2A and 2B. Both the synchronous DRAMs 2A and 2B have the same configuration as the synchronous DRAM 2. The synchronous DRAM 2A, one of them, is set up with a burst length of 16 bytes, whereas the synchronous DRAM 2B, the other of them, is set up with a burst length of 32 bytes. The CPU 3 independently sets the burst length of each of the synchronous DRAMs 2A and 2B in the mode registers of the synchronous DRAMs 2A and 2B by software after power on reset. At this time, the memory control registers 23 inside the memory control part 6 are set up with bus control information for a burst length of external memories such as the synchronous DRAMs 2A and 2B. The other configurations are the same as Fig. 1, omitting the detailed description.

According to the data processing system shown in Fig. 10, when a burst length (16 bytes, for example) relatively shorter than the cache line length (32 bytes, for example) of the cache memory 4 is set in the synchronous DRAM 2A, a plurality of block data outputted by wraparound from the synchronous DRAM 2A can be joined and filled in the cache memory 4. When the synchronous DRAM 2B set up with the burst length equal to the cache line length is included in the data processing system, the cache fill operation corresponding to the burst length of the synchronous DRAM 2B is also allowed in processing a cache mishit for the synchronous DRAM 2B.

Additionally, in the case where write through is adopted

as the process for a cache write hit of the cache memory 4, when write data is written into the synchronous DRAM 2A from the write through buffer shorter than the cache line length (eight bytes, for example), there are fewer unnecessary data transfer cycles because of the relatively shorter burst length. For 12 bytes in the latter half of the burst access operation at this time, the data masking signal DM can mask data to suppress the actual data write operation. When the synchronous DRAM 2B set up with the burst length equal to the cache line length is the object to be written by write through, unnecessary cycles are increased as compared with the synchronous DRAM 2A even when performing write mask. However, the amount of data to be accessed or transferred to the synchronous DRAM 2B at one time can be increased in a state that the synchronous DRAM 2B is temporarily removed from the object for cache, and thus the data processing system can contribute to the enhanced data processing performance. The control of temporarily removing the synchronous DRAM 2B from the object for cache can be performed by operation modes of the microprocessor 1 or by the CPU 3 to set the cache control register in the cache control part 5, not shown.

Accordingly, in the data processing system where a different burst length is set in a plurality of synchronous DRAMs, it can suppress unnecessary cycles as much as possible in transferring data of eight bytes, relatively small, without

deteriorating the performance of efficiently transferring high capacity data of 32 bytes, and it can realize various connection configurations or utilization forms of a plurality of memories having a different burst length.

Furthermore, in the configuration of programs operated in the microprocessor 1, the synchronous DRAM 2B is allowed to hold data having a size equal to or greater than a program cord and a cache line length, and the synchronous DRAM 2A is allowed to hold data having a size smaller than a cache line length, whereby the enhanced processing performance of the microprocessor 1 can be intended as well.

The invention made by the inventors has been described in detail in accordance with the embodiments, but the invention is not limited thereto. It is needless to say that the invention can be modified variously within the scope of the teachings, not deviating it.

For example, the cache memory may be for storing programs, or for storing data and programs as mixed. Additionally, for the cache memory, the associative memory format such as set associative, full associative or direct map can be adopted. Furthermore, for the cache memory, the write back method may be adopted instead of the write through method.

Moreover, the data processing device may be integrated with other computing units such as a floating point processing unit, other bus master modules such as a direct memory access

controller, or other peripheral circuits such as a timer or RAM. The burstable memory is not defined to the synchronous DRAM; it may be a synchronous SRAM. The number of burstable memories included in the data processing system may be increased properly.

The invention can be widely adapted to a data processing device capable of having access to a burstable memory and the data processing system. For example, it can be adapted to various data processing devices formed with a semiconductor integrated circuit such as a microprocessor, a microcomputer, a data processor, and a DSP.

20050213-030502